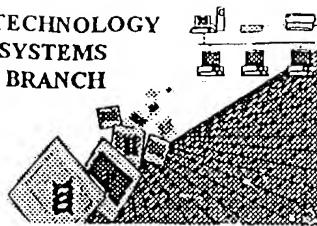


BIOTECHNOLOGY
SYSTEMS
BRANCH



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 101761,006
Source: EFSWD
Date Processed by STIC: 11/30/04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT
MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>), EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

Raw Sequence Listing Error Summary

ERROR DETECTED SUGGESTED CORRECTION SERIAL NUMBER: 101761,006

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleic
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220> Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 00/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n/Xaa "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



IFWO

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/761,006

DATE: 01/30/2004
 TIME: 14:36:38

Input Set : A:\seqlist.txt
 Output Set: N:\CRF4\01302004\J761006.raw

SEQUENCE LISTING

- 4 (1) GENERAL INFORMATION:
 - 6 (i) APPLICANT: Oon, Chong Jin
 - 7 Lim, Gek Keow
 - 8 Zhao, Yi
 - 9 Chen, Wei Ning
 - 11 (ii) TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
 - 12 USES THEREOF
 - 14 (iii) NUMBER OF SEQUENCES: 11
 - 16 (iv) CORRESPONDENCE ADDRESS:
 - 17 (A) ADDRESSEE: Ladas & Parry
 - 18 (B) STREET: 26 West 61 Street
 - 19 (C) CITY: New York
 - 20 (D) STATE: New York
 - 21 (E) COUNTRY: USA
 - 22 (F) ZIP: 10023
 - 24 (v) COMPUTER READABLE FORM:
 - 25 (A) MEDIUM TYPE: Floppy disk
 - 26 (B) COMPUTER: IBM PC compatible
 - 27 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - 28 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 - 30 (vi) CURRENT APPLICATION DATA:
 - C--> 31 (A) APPLICATION NUMBER: US/10/761,006
 - C--> 32 (B) FILING DATE: 20-Jan-2004
 - 33 (C) CLASSIFICATION: 435
 - 35 (vii) PRIOR APPLICATION DATA:
 - 36 (A) APPLICATION NUMBER: PCT/SG98/00046
 - 37 (B) FILING DATE: 19-JAN-1998
 - 39 (viii) ATTORNEY/AGENT INFORMATION:
 - 40 (A) NAME: Mass, Clifford J.
 - 41 (B) REGISTRATION NUMBER: 30,086
 - 42 (C) REFERENCE/DOCKET NUMBER: U-014987-0
 - 44 (ix) TELECOMMUNICATION INFORMATION:
 - 45 (A) TELEPHONE: (212) 708-1800

*Does Not Comply
 Corrected Diskette Needed
 (P5.Q-S)*

ERRORED SEQUENCES

- 48 (2) INFORMATION FOR SEQ ID NO: 1:
 - 50 (i) SEQUENCE CHARACTERISTICS:
 - 51 (A) LENGTH: 3215 base pairs
 - 52 (B) TYPE: nucleic acid
 - 53 (C) STRANDEDNESS: double

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/761,006

DATE: 01/30/2004
TIME: 14:36:38

Input Set : A:\seqlist.txt
Output Set: N:\CRF4\01302004\J761006.raw

54 (D) TOPOLOGY: circular
 58 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 E--> 60 CTCCACAACA TTCCACCAAG CTCTGCTAGA TCCCAGGGTG AGGGGCCTAT
 61 ATTTTCCTGC 60
 E--> 63 TGGTGGCTCC AGTTCCGGAA CAGTAAACCC TGTTCCGACT ACTGCCTCTC
 64 CCATATCGTC 120
 E--> 66 AATCTTCTCG AGGACTGGGG ACCCTGCACC GAACATGGAG AACACAACAT
 67 CAGGATTCCCT 180
 E--> 69 AGGACCCCTG CTCGTGTTAC AGGCAGGGTT TTTCTCGTTG ACAAGAATCC
 70 TCACAATACC 240
 E--> 72 GCAGAGTCTA GACTCTGGTG GACTTCTCTC AATTTCTAG GGGGAGCACC
 73 CACGTGTTCC 300
 E--> 75 TGGCCAAAAT TCGCAGTCCC CAACCTCCAA TCACTCACCA ACCTCTTGTG
 76 CTCCAATTG 360
 E--> 78 TCCTGGCTAT CGCTGGATGT GTCTGCGGGG TTTTATCATA TTCCTCTTCA
 79 TCCTGCTGCT 420
 E--> 81 ATGCCTCATC TTCTTGTGG TTCTTCTGGA CTACCAAGGT ATGTTGCCCG
 82 TTGTCCTCT 480
 E--> 84 ACTTCCAGGA ACATCAACCA CCAGCACGGG GCCATGCAAG ACCTGCACGA
 85 CTCCGTCTCA 540
 E--> 87 AGGAAAATCT ACGTTCCCT CTTGTTGCTG TACAAAACCT TCGGACGGAA
 88 ACTGCACTTG 600
 E--> 90 TATTCCCATC CCATCATCCT GGGCTTCGC AAGATTCCTA TGGGAGTGGG
 91 CCTCAGTCGG 660
 E--> 93 TTCTCCTGG CTCAGTTTAC TAGTGCCATT TGTTCAAGTGG TTCGTAGGGC
 94 TTCCCCCAC 720
 E--> 96 TGTTTGGCTT TCAGTTATAT GGATGATGTG GTATTGGGG CGAAGTCTGT
 97 ACAACATCTT 780
 E--> 99 GAGTCCCTTT TTACCTCTAT TACCAATTCTT CTTTGTCTT TGGGTATACA
 100 TTAAACCTT 840
 E--> 102 AATAAAACCA AACGTTGGGG CTACTCCCTT AACTTCATGG GATATGTAAT
 103 TGGAAAGTTGG 900
 E--> 105 GGTACTTTAC CGCAGGAACA TATTGTACTA AAACCTCAAGC AATGTTTCG
 106 AAAACTGCCT 960
 E--> 108 GTAAATAGAC CTATTGATTG GAAAGTATGT CAAAGAATTG TGGGTCTTT
 109 GGGCTTGCT 1020
 E--> 111 CCCCTTTTA CACAATGTGG CTATCCTGCC TTGATGCCCTT TATATGCATG
 112 TATACAATCT 1080
 E--> 114 AAGCAGGCTT TCACTTTCTC GCCAACTTAC AAGGCCTTTC TGTGTAAACA
 115 ATATCTGAAC 1140
 E--> 117 CTTTACCCCCG TTGCCCGCA ACGGTCCGGT CTCTGCCAAG TGTTGCTGA
 118 CGCAACCCCC 1200
 E--> 120 ACTGGATGGGG GCTTGGCCAT AGGCCATCAG CGCATGGCTG GAACCTTCT
 121 GGCTCCTCTG 1260
 E--> 123 CCGATCCATA CTGCGGAACCT CCTAGCAGCT TGTTTGCTC GCAGCCGGTC
 124 TGGAGCAAA 1320
 E--> 126 CTTATCGGAA CCGACAACTC TGTTGTCCCTC TCTCGGAAAT ACACCTCCTT
 127 TCCATGGCTG 1380
 E--> 129 CTAGGGTGTG CTGCCAACTG GATCCTGCCG GGGACGTCTT TTGTCTACGT

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/761,006

DATE: 01/30/2004
TIME: 14:36:38

Input Set : A:\seqlist.txt
Output Set: N:\CRF4\01302004\J761006.raw

130	CCCGTCGGCG	1440			
E--> 132	CTGAATCCCG	CGGACGACCC	GTCTCGGGGC	CGTTGGGGC	TCTACCGTCC
133	CCTTCTTCAT	1500			
E--> 135	CTGCCGTTCC	GGCCGACCAC	GGGGCGCAC	TCTCTTACG	CGGTCTCCCC
136	GTATGTGCCT	1560			
E--> 138	TCTCATCTGC	CGGACCGTGT	GCACTTCGCT	TCACCTCTGC	ACGTCGCATG
139	GAGACCACCG	1620			
E--> 141	TGAACGCAAG	CCAGGTCTTG	CCCAAGGTCT	TATATAAGAG	GACTCTTGGA
142	CTCTCAGCAA	1680			
E--> 144	TGTCAACGAC	CGACCTTGAG	GCATACTTCA	AAGACTGTGT	GTTTAAAGAC
145	TGGGAGGAGT	1740			
E--> 147	TGGGGGAGGA	GATTAGGTTA	AAGATTATG	TACTAGGAGG	CTGTAGGCAT
148	AAATTGGTCT	1800			
E--> 150	GTTCACCAGC	ACCATGCAAC	TTTTTCTCCT	CTGCCTAATC	ATCTCATGTT
151	CATGTCCTAC	1860			
E--> 153	TGTTCAAGCC	TCCAAGCTGT	GCCTTGGGTG	GCTTTGGGAC	ATGGACATTG
154	ACCCGTATAA	1920			
E--> 156	AGAATTGGA	GCATCTGCTG	AGTTACTCTC	TTTTTGCCCT	TCTGACTTCT
157	TTCCGTCTAT	1980			
E--> 159	TCGAGATCTC	CTCGACACCG	CCTCTGCTCT	GTATCGGGAG	GCCTTAGAGT
160	CTCCGGAACA	2040			
E--> 162	TGTTCGCCT	CACCATACAG	CACTCAGGCA	AGCTATTTG	TGTTGGGTG
163	AGTTGATGAA	2100			
E--> 165	TCTGGCCACC	TGGGTGGGAA	GTAATTGGA	AGATCCAGCA	TCCAGGGAAT
166	TAGTAGTCAG	2160			
E--> 168	CTATGTCAC	GTAAATATGG	GCCTAAAAC	CAGACAAATA	TTGTGGTTTC
169	ACATTTCTG	2220			
E--> 171	TCTTACTTT	GGAAGAGAAA	CTGTTCTGA	GTACTTGGTA	TCTTTGGAG
172	TGTGGATTG	2280			
E--> 174	CACTCCTACC	GCTTACAGAC	CACCAAATGC	CCCTATCTTA	TCAACACTTC
175	CGGAAACTAC	2340			
E--> 177	TGTTGTTAGA	CGACGAGGCA	GGTCCCCTAG	AAGAAGAACT	CCCTCGCCTC
178	GCAGACGAAG	2400			
E--> 180	GTCTCAATCG	CCGCGTCGCA	GAAGATCTCA	ATCTCGGGAA	TCTCAACGTT
181	AGTATTCTT	2460			
E--> 183	GGACTCATAA	GGTGGGAAAC	TTTACTGGGC	TTTATTCTTC	TACTGTACCT
184	GTCTTTAAC	2520			
E--> 186	CCGAGTGGCA	AATTCCCTCC	TTTCCTCACA	TTCATTTACA	AGAGGACATT
187	ATTAATAGAT	2580			
E--> 189	GTCAACAAATA	TGTGGGCCCT	CTTACAGTTA	ATGAAAAAAG	AAGATTAAAA
190	TTAATTATGC	2640			
E--> 192	CTGCTAGGTT	TTATCCTAAC	CTTACTAAAT	ATTTGCCCTT	AGACAAAGGC
193	ATTAACCGT	2700			
E--> 195	ATTATCCTGA	ACATGCAGTT	AATCATTACT	TCAAAACTAG	GCATTATTAA
196	CATACTCTGT	2760			
E--> 198	GGAAGGCTGG	CATTCTATAT	AAGAGAGAAA	CTACACGCAG	CGCCTCATT
199	TGTGGGTAC	2820			
E--> 201	CATATTCTTG	GGAACAAAGAG	CTACAGCATG	GGAGGTTGGT	CTTCCAAACC
202	TCGACAAGGC	2880			

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/761,006

DATE: 01/30/2004
TIME: 14:36:38

Input Set : A:\seqlist.txt
Output Set: N:\CRF4\01302004\J761006.raw

E--> 204 ATGGGGAGCA ATCTTGCTGT TCCCAATCCT CTGGGATTCT TTCCCGATCA
205 CCAGTTGGAC 2940
E--> 207 CCTGCGTTCG GAGCCAACTC AAACAATCCA GATTGGGACT TCAACCCAA
208 CAAGGATCAC 3000
E--> 210 TGGCCAGAGG CAAATCAGGT AGGAGTGGGA GCATTGGGC CAGGGTTCAC
211 CCCACCACAC 3060
E--> 213 GGCGGTCTTT TGGGGGGGAG CCCTCAGGCT CAGGGCATAT TGACAACAGT
214 GCCAGCAGCA 3120
E--> 216 CCTCCTCCTG CCTCCACCAA TCGGCAGTCA GGAAGACAGC CTACTCCCAC
217 CTCTCCACCT 3180
219 CTAAGAGACA GTCATCCTCA GGCCACGCAG TGGAA 3215

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/761,006

DATE: 01/30/2004

TIME: 14:36:39

Input Set : A:\seqlist.txt

Output Set: N:\CRF4\01302004\J761006.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:60 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:1
M:254 Repeated in SeqNo=1

The type of errors shown exist throughout
the Sequence Listing. Please check subsequent
sequences for similar errors.